MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Switchgear Partial Discharge Diagnosis Using Scarce Fault Records

Sun, Hongbo; Otake, Yasutomo; Matsuyama, Kotaro; Raghunathan, Arvind

TR2025-155 October 23, 2025

Abstract

SF6 gas has traditionally been used in Gas- Insulated Switchgear, but due to its extremely high Global Warming Potential, there is growing interest in alternative insulating media, such as Green Gas for Grid and dry air. As a result, there is an increasing need to develop new partial discharge (PD) diagnostic methods tailored to these alternative media, while also addressing the challenge of limited fault data. In this paper, using high-pressure dry air as an example, we propose a methodology for adapting existing PD diagnostic models—originally developed for atmospheric conditions—to high-pressure dry air, leveraging transfer learning techniques. The proposed method first transforms the raw data of measured applied voltages and partial discharge voltages into six relevant features for each time window through feature engineering. These features are then fed into the PD pattern diagnosis model, which consists of a feature extractor, a PD fault type classifier, and a domain discriminator. Feature discrepancy loss, including maximum mean discrepancy, batch-based instance separation, and batch-based feature decorrelation, is added to the loss function to optimize the model's parameters. We evaluate the prediction performance under varying levels of data scarcity for high-pressure dry air switchgear. Additionally, we compare the estimation performances of transfer learning versus deep learning and discuss the transition point between these two approaches as the dataset evolves.

IEEE PES Innovative Smart Grid Technologies Conference - Europe (ISGT Europe) 2025

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Switchgear Partial Discharge Diagnosis Using Scarce Fault Records

Hongbo Sun
Optimization and Intelligent Robotics
Mitsubishi Electric Research Laboratories
Cambridge, MA 02139, US
hongbosun@merl.com

Yasutomo Otake
Advanced Technology R&D Center
Mitsubishi Electric Corporation
Amagasaki 661-8661, Japan
Otake.Yasutomo@ce.MitsubishiElectric.co.jp

Kotaro Matsuyama
Advanced Technology R&D Center
Mitsubishi Electric Corporation
Amagasaki 661-8661, Japan
Matsuyama.Kotaro@dn.MitsubishiElectric.co.jp

Arvind Raghunathan
Optimization and Intelligent Robotics
Mitsubishi Electric Research Laboratories
Cambridge, MA 02139, USA
raghunathan@merl.com

***Abstract*—SF_6 gas has traditionally been used in Gas-Insulated Switchgear, but due to its extremely high Global Warming Potential, there is growing interest in alternative insulating media, such as Green Gas for Grid and dry air. As a result, there is an increasing need to develop new partial discharge (PD) diagnostic methods tailored to these alternative media, while also addressing the challenge of limited fault data. In this paper, using high-pressure dry air as an example, we propose a methodology for adapting existing PD diagnostic models—originally developed for atmospheric conditions—to high-pressure dry air, leveraging transfer learning techniques. The proposed method first transforms the raw data of measured applied voltages and partial discharge voltages into six relevant features for each time window through feature engineering. These features are then fed into the PD pattern diagnosis model, which consists of a feature extractor, a PD fault type classifier, and a domain discriminator. Feature discrepancy loss, including maximum mean discrepancy, batch-based instance separation, and batch-based feature decorrelation, is added to the loss function to optimize the model's parameters. We evaluate the prediction performance under varying levels of data scarcity for high-pressure dry air switchgear. Additionally, we compare the estimation performances of transfer learning versus deep learning and discuss the transition point between these two approaches as the dataset evolves.**

Keywords— deep learning, partial discharge, scare fault data, switchgear, transfer learning

I. Introduction

The stability and safety of electrical power systems heavily depend on the health of electrical power equipment. Electrical insulation plays a critical role in the health of high-voltage equipment, as insulation failure accounts for over 60% of high-voltage equipment failures [1]. When insulation fails, partial discharge (PD) occurs. Therefore, monitoring PD in power equipment is essential to prevent significant failures and power outages.

According to IEC 60270 [2], a partial discharge (PD) is defined as "a localized electrical discharge that partially bridges the insulation between conductors, which may or may not occur adjacent to a conductor." When a PD occurs, it generates a rapid transient pulse. This signal is closely associated with the PD source, and different types of PD defects produce distinct signal patterns. Identifying the PD source helps determine the condition of the equipment.

Switchgear, a critical component of electrical power systems, plays a vital role in ensuring system integrity and reliability. Utilizing partial discharge monitoring to assess the health of switchgear is essential for effective maintenance and operational oversight. However, building a reliable PD diagnostic model requires a large volume of fault data to achieve acceptable accuracy. In practice, obtaining sufficient data for newly developed switchgear types can be challenging. For example, SF_6 gas has traditionally been used in Gas-Insulated Switchgear, but due to its extremely high Global Warming Potential, there is increasing interest in alternative insulating gases such as Green Gas for Grid or dry air. As a result, there is a need to develop new PD diagnostic methods specifically designed for these alternative insulation materials, while also addressing the challenge of limited fault data availability.

Various approaches [3]~[5] have been proposed to monitor, detect, and diagnose partial discharge (PD) in switchgears. These approaches can generally be classified into two categories: model-based approaches [7],[14]~[16] and data-driven or model-free approaches. Model-based approaches aim to develop mathematical models to predict and track the degradation progression of switchgears. However, constructing models with an appropriate level of complexity is challenging, as the mechanical principles and degradation mechanisms of switchgears are often complex and not fully understood.. On the other hand, data-driven approaches [6]~[13] provide a straightforward solution by utilizing large volumes of historical data to infer PD fault modes without requiring prior theoretical knowledge. However, these methods demand high-quality, abundant training data, which presents a significant challenge in practical applications.

In recent years, model-free machine learning (ML) techniques such as deep learning [10]~[11], transfer learning [12]~[13], which do not rely on a predefined parametric model, have shown promising improvements across a wide range of applications. Particularly, machine learning's ability to uncover complex, hidden patterns in data has proven highly successful, often outperforming state-of-the-art human-designed algorithms. These advancements have provided valuable tools for classifying PD fault patterns in switchgears. However, machine learning-based approaches still face several challenges, including: (1) insufficient training data, (2) unclear data representation, and (3) the need to adapt neural network architectures for different types of switchgears.

Therefore, there is a need to develop more advanced methods for diagnosing partial discharge in switchgears that offer improved generalization capabilities.

This paper proposes a transfer learning-based method for diagnosing partial discharge faults in switchgear. The measured fault signals from the switchgear, including applied voltages and partial discharge voltages, are first denoised using discrete Fourier transform and discrete wavelet transform techniques. Next, feature engineering is employed to generate multiple characteristic features based on the statistical distribution of the switchgear measurements using a sliding window algorithm. Six different statistic-based features are generated, including the average value of applied voltage to represent the applied voltage magnitude and event timing, and the mean, standard deviation, Kurtosis, Skewness, and total number of spikes of partial discharge voltages to represent the variation of partial discharge within the time block. This approach represents the features of partial discharges with a reduced data volume, without losing the original time characteristics of partial discharge faults in the switchgear. The partial discharge pattern classification model consists of a feature extractor, a PD fault type classifier, and a domain discriminator. The parameters of the pattern classification model are optimized by combining classifier loss, domain discrepancy loss, and feature discrepancy loss. The feature loss includes maximum mean discrepancy loss, batch-based instance separation loss, and batch-based feature decorrelation loss.

The remainder of this paper is organized as follows: Section II presents the PD defect models and experiments. Section III describes the proposed transfer learning-based PD defect type detection method. Sections IV and V provide the results obtained using the transfer learning based method, the deep learning based method, and the benchmark Support Vector Machine method. The conclusion is drawn in Section VI.

II. PD Defect Models and Experiments

Fig. 1 illustrates the PD defect types simulated in a laboratory switchgear insulated with atmospheric or high-pressure air. These defect types include floating electrodes, metal protrusions, metal protrusions into the insulator, and surface discharge. The atmospheric conditions for the PD test are maintained at a temperature of 20–23°C, humidity of 50–60%, and pressure ranging from 1009 hPa to 1020 hPa. Additionally, high-pressure dry air is maintained at 0.9 MPa (absolute).

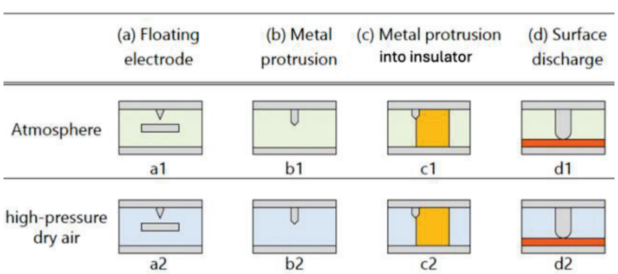


Fig 1. PD Defect Models with Different Insulation Mediums

Fig. 2 illustrates the experimental setup for PD testing and data collection using IEC 60270 PD detection method [2]. Electrical measurements are captured on an oscilloscope from the laboratory switchgear during PD fault simulation. The measurement data are obtained from the laboratory model under two different insulation conditions: atmospheric air and high-pressure dry air. These measurements are collected from two oscilloscope channels: the first channel captures the AC voltage waveform, while the second channel records the PD signal. Both the applied voltage signal from the first channel and the PD signal from the second channel are used for diagnostic purposes.

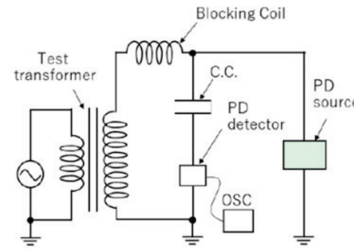


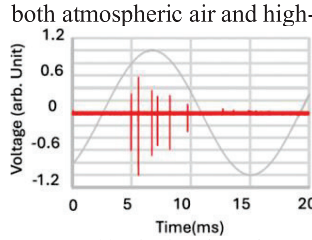
Fig 2. Experiment setup for using IEC 60270 PD detection method

Table I summarizes the collected datasets for different defect types used in this paper. In total, 6,194 and 7,427 fault events were recorded for atmosphere-based and high-pressure air-based switchgear systems, respectively. For all events—except for the floating electrode defect in the atmosphere-based switchgear—2,000 data points were collected over 0.02 seconds (sampling frequency: 100 kHz; time resolution: 10 μs). In the exceptional case of the floating electrode defect, 4,000 data points were collected over the same duration (sampling frequency: 200 kHz; time resolution: 5 μs), corresponding to a single event (i.e., one sample).

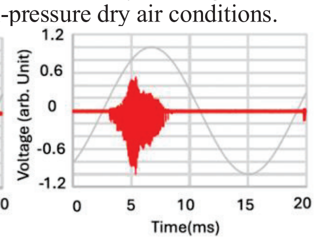
TABLE I. PD Data Summaries

Defect Type	Total number of Samples	
	Atmosphere based switchgear	***Air based Switchgear***
(a). Floating electrode	1769	1626
(b). Metal protrusion	1313	1335
(c). Metal protrusion into insulator	1788	2683
(d).Surface discharge	1324	1783
Totals	6194	7427

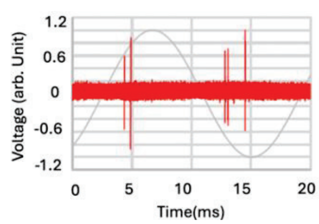
Figs 3 and 4 illustrate examples of different defect types observed in the applied voltage and partial discharge measurements collected from the laboratory switchgear under both atmospheric air and high-pressure dry air conditions.



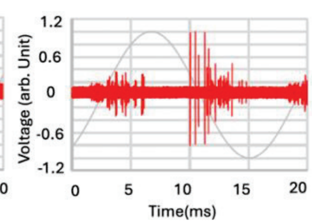
(a). Floating electrode.



(b) Metal protrusion

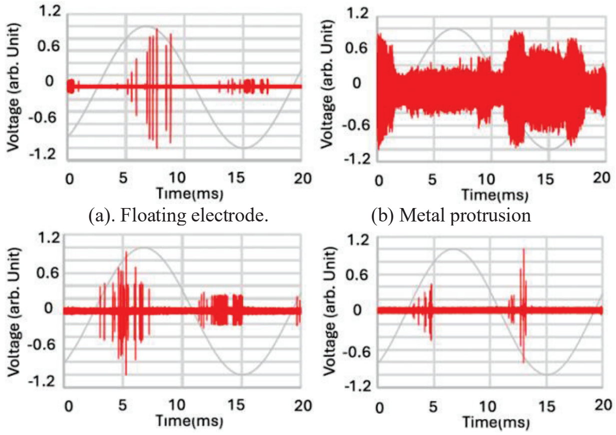


(c). Metal protrusion into insulator



(d) Surface discharge

Fig 3. Exemplar samples for PD faults in an atmosphere based switchgear



(a). Floating electrode. (b) Metal protrusion

(c). Metal protrusion into Insulator. (d) Surface discharge

Fig 4. Exemplar samples for PD faults in a high-pressure dry air based switchgear

In the figures, the gray curves represent the ideal applied voltages, while the red curves represent the denoised partial discharge voltages. It is evident that the PD waveforms differ significantly both between different defect types under the same insulation medium and between the same defect type under different insulation mediums. Directly differentiating defect types using waveforms is challenging. Instead, the statistics of the PD waveforms, such as their moments and the number of spikes (both positive and negative spikes included), are used to represent these events.

III. Transfer Learning based PD Defect Type Detection

Fig. 5 illustrates the data preprocessing procedure applied before feeding the processed labeled datasets for partial discharge (PD) defect type detection.

The raw data is first denoised using the Discrete Fourier Transform (DFT) technique to filter the applied voltage signals, and the Discrete Wavelet Transform (DWT) technique to filter the PD signals. After normalization, the labeled dataset undergoes feature engineering, where one feature is derived from the applied voltages, and five features are extracted from the partial discharge voltages for each time block. The sliding window method is used to divide the entire measurement period into overlapping time blocks of equal width. In this paper, each original sample event is transformed into WWW block data samples (e.g., W=99W = 99W=99). The block width of each window is denoted as SWSWSW (e.g., 0.0004 s), and the gap between two consecutive windows is denoted as GWGWGW (e.g., 0.0002 s).

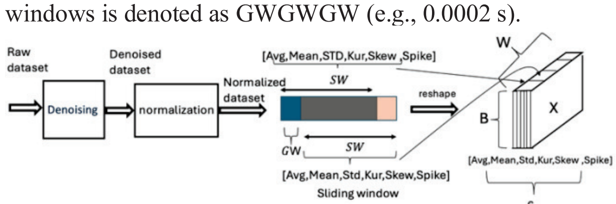


Fig 5. Sliding window based data preprocessing

For each block data sample, 6 different statistic-based features are calculated, including the average value of applied voltage (Avg) to represent the applied voltage magnitude and event timing, and the mean (Mean), standard deviation (Std), Kurtosis (Kur), Skewness (Skew), and total number of spikes (Spike) of partial discharge voltages to represent the variation of partial discharge within the time block. Thus, a series of time-windowed features are generated. After that, reshaping is applied to generate input tensor X_1, for the source domain i.e. atmosphere switchgear, whose size of $B \times 6 \times W$, where B denotes the batch size. Similarly, input tensor X_2 for the target domain i.e. air switchgear can be generated, whose size of $B \times 6 \times W$. The batch can be generated by randomly selecting samples from available collected sample events from the switchgear with given number of times, i.e. the size of batch.

Fig. 6 shows the process for transfer learning based partial discharge pattern classification. The process uses a feature extractor to extract representative features to represent fault events, a PD pattern classifier to classify the fault type based on the extracted representative features, a domain discriminator being used for domain adaptation (DA). The pattern classification process first forms a concatenated tensor, $X = [X_1 : X_2]$, and then input to the feature extractor F , $F = [F_1 : F_2]$, an adversarial net based domain discriminator being used for domain adaptation, and PD fault classifier, and relative feature loss computation. F_1 and F_2 are the corresponding extracted representative features from the input features X_1 and X_2 for the source and target switchgear, respectively.

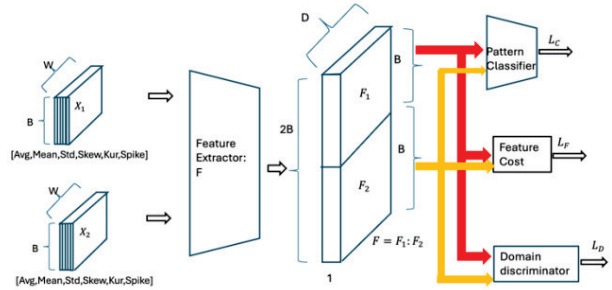


Fig 6. Transfer learning based PD diagnosis framework

The model's parameters are optimized using a loss function that is defined as the sum of the classifier loss L_C (using cross-entropy loss), the domain adaptation loss L_D (also using cross-entropy loss), and the feature loss L_F:

$$minimize\ Loss = L_C + L_D + L_F. \tag{1}$$

Feature loss is used to balance between feature selectivity among different features, and feature invariance among source and target domains. It is defined as a weighed sum of maximum mean discrepancy loss $\mathrm{MMD}(F_1, F_2)$, batch-based instance separation loss [17] $\mathrm{BIS}(F_1, F_2)$, and batch-based feature decorrelation loss [17] $\mathrm{BFD}(F_1^T, F_2^T)$:

$$L_F = \lambda_{MMD}\,\mathrm{MMD}(F_1, F_2) + \lambda_{BIS}\mathrm{BIS}(F_1, F_2) + \lambda_{BFD}\mathrm{BFD}(F_1^T, F_2^T), \tag{2}$$

where λ_{MMD}, λ_{BIS} and λ_{BFD} denote the weighting coefficients for the discrepancy cost, the instance separation cost, and the feature decorrelation loss, respectively.

The batch-based instance separation loss, $\mathrm{BIS}(F_1, F_2)$ is to encourage the network to learn different features for each training example, and defined as:

$$\mathrm{BIS}(F_1, F_2) = -\frac{1}{2 \times B \times D}\sum_{i=1}^{B} \log\left(\frac{\exp(F_{1i}(F_{1i})^T)}{\sum_{j=1}^{B} \exp\left(F_{1i}\left(F_{1j}\right)^T\right)}\right) - \frac{1}{2 \times B \times D}\sum_{i=1}^{B} \log\left(\frac{\exp(F_{2i}(F_{2i})^T)}{\sum_{j=1}^{B} \exp\left(F_{2i}\left(F_j\right)^T\right)}\right) \tag{3}$$

where D is the feature size of feature exactor bottleneck layer. B denotes the batch size. W is the number of sliding windows. F_{1i} and F_{2i} denote the i-th row vector of $F_1$and F_2. The batch-based feature decorrelation loss is encourages the network to learn distinct features, and .defined as:

$$\mathrm{BFD}(F_1, F_2) = -\frac{1}{2 \times B \times D} \sum_{i=1}^{D} \log\left(\frac{\exp(G_{1i}(G_{1i})^T)}{\sum_{j=1}^{D} \exp\left(G_{1i}(G_{1j})^T\right)}\right) - \frac{1}{2 \times B \times D} \sum_{i=1}^{D} \log\left(\frac{\exp(G_{2i}(G_{2i})^T)}{\sum_{j=1}^{D} \exp\left(G_{2i}(G_{2j})^T\right)}\right) \quad (4)$$

where $G_1 = (F_1)^T, G_2 = (F_2)^T$. G_{1i} and G_{2i} denote the i-th row vector of $G_1$and G_2.

The AdamW algorithm [18] is used as the optimizer for learning. It is a stochastic gradient descent method that is based on adaptive estimation of first-order and second-order moments with an added method to decay weights. The learning rate is usually set as 0.01.

Fig. 7 illustrates the detailed configuration of the different layers in the feature extractor, the domain discriminator, and the pattern classifier. Conv1D, ReLU, Pool, Dropout, and Linear transformer represent a 1D convolutional layer, a rectified linear unit, a max pooling layer, a dropout layer, and a fully connected layer, respectively. The kernel size and padding of the Conv1D is set as 3 and 1. The kernel size, stride, and padding of the Pool are set as 2,2 and 0. The drop-out is set as 0.5.

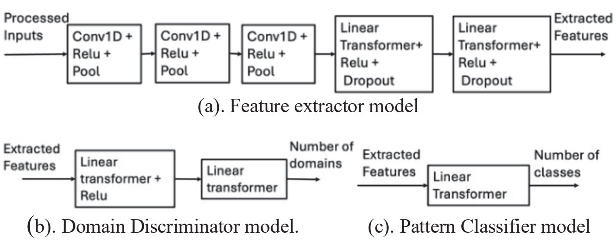


Fig 7. Components of PD diagnosis framework

IV. PD Defect Type Detection with Scare Target Datasets

To evaluate the effectiveness of domain adaptation (i.e. transfer learning), we have created a set of test scenarios with varying levels of data scarcity as shown in Table II. In our experiments, we intentionally reduce the amount of training data from the target domain to simulate varying levels of data scarcity for the target type of switchgears, while keeping the amount of training data from the source domain unchanged.

TABLE II. Target Data Scarity Scenarios

Scenario #	**Training samples**				
	Atmosphere based switchgear		***Air based Switchgear***		
	Percentage	***Total number of samples***	***Total defect types***	***Percentag e***	***Total number of samples***
1	80%	4955	4	80%	5941
2	80%	4955	4	10%	743
3	80%	4955	4	1%	74
4	80%	4955	4	0.5%	37
5	80%	4955	4	0.1%	7
6	80%	4955	1	0.014%	1

Table III presents the estimation performance metrics, including accuracy, precision, recall, and F1-score, under varying levels of target data scarcity. To mitigate the effect of class imbalance among defect types, the macro-averages of precision, recall, and F1-score are reported. For performance evaluation, the training dataset is randomly sampled from the total target domain data, while the remaining data are reserved for testing. The training process is performed for 1000 epochs. As shown in Table III, the estimation performance improves as the number of target training samples increases, achieving an accuracy of 94.96% when the training set comprises 743 samples.

TABLE III. Estimation Performance Metrics for Transfer Learning

Total Number of Samples	**Accuracy**	**Macro-averaged Precision**	**Macro-averaged Recall**	**Macro-averaged F1-score**
5941	0.9596	0.9587	0.9524	0.9535
743	0.9496	0.9580	0.9548	0.9554
74	0.8728	0.8867	0.8913	0.8847
37	0.8429	0.8725	0.8613	0.8548
7	0.7586	0.8056	0.7799	0.7525
1	0.5990	0.6919	0.6120	0.5493

V. PD Defect Type Identification Using Deep Learning vs. Transfer Learning

Fig. 8 illustrates the deep learning-based PD diagnosis framework, which is configured with a feature extractor and a pattern classifier, both using the same layer configurations defined in Fig. 7. The model's parameters are optimized using the classifier loss L_C (using cross-entropy loss):

$$minimize\ Loss = L_C. \quad (5)$$

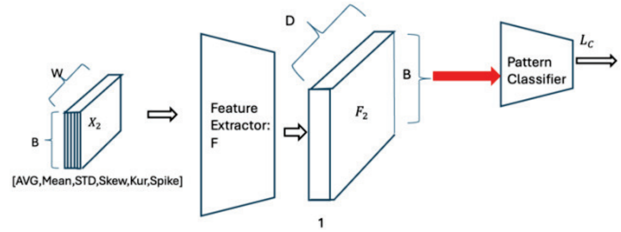


Fig 8. Deep learning based PD diagnosis framework

Similarly, we simulate different levels of target data scarcity for air-based switchgear by intentionally reducing the amount of target domain training data, without using the source domain datasets.

Table IV lists the estimation performance metrics achieved using deep learning, based solely on the target data. Comparing the results obtained using deep learning, as shown in Table IV, with those obtained through transfer learning, as shown in Table III, we verified that transfer learning can improve prediction performance by learning from source domain datasets when the target domain data is significantly scarce (i.e., the total number of target samples is less than 37). However, once the target domain accumulates enough data (i.e., the total number of target samples exceeds 37), its prediction performance decreases when domain adaptation is used to mix its characteristics with those of the source domain.

For performance comparison, we employ a linear Support Vector Machine (SVM) as a benchmark. The inputs to the SVM are aligned with those of the feature extractor in Fig. 8, while the outputs correspond to those of the pattern classifier in Fig. 8. Table V presents the estimation performance metrics achieved by the SVM, evaluated solely on the target data.

TABLE IV. ESTIMATION PERFORMANCE METRICS FOR DEEP LEARNING

Total Number of Samples	Accuracy	Macro-averaged Precision	Macro-averaged Recall	Macro-averaged f1-score
5941	0.9966	0.9964	0.9959	0.9961
743	0.9864	0.9870	0.9822	0.9843
74	0.9202	0.9470	0.9420	0.9408
37	0.8429	0.8930	0.8259	0.8383
7	0.6171	0.5644	0.6008	0.5477
1	0.2400	0.0596	0.2500	0.0962

TABLE V. ESTIMATION PERFORAMCE METRICS FOR SVM

Total Number of Samples	Accuracy	Macro-averaged Precision	Macro-averaged Recall	Macro-averaged f1-score
5941	0.9879	0.9873	0.9888	0.9880
743	0.9708	0.9787	0.9716	0.9747
74	0.8650	0.8958	0.8617	0.8712
37	0.8166	0.8483	0.8034	0.8127
7	0.6377	0.5386	0.6064	0.5437
1	0.2607	0.3115	0.2765	0.1467

Comparing the results in the above tables, the best estimation performance can be achieved by dynamically switching between transfer learning and deep learning (or SVM) as the target data volume evolves. Taking transfer learning and deep learning as examples, Table VI shows that transfer learning is preferable when the target dataset is small, whereas deep learning becomes more effective once the dataset size surpasses a critical point. This critical point can be identified when the dataset includes all defect types and the number of samples for each defect type exceeds an empirical threshold—typically around 20 samples. This threshold can be established by analyzing model performance across varying sample sizes using a well-labeled dataset collected from similar types of switchgear. The minimum sample count at which the model achieves acceptable performance can then be adopted as the practical threshold.

TABLE VI. SWITCHING BETWEEN TRANSFER LEARNING AND DEEP LEARNING

Total Number of Samples	Estimation accuracy using transfer learning	Estimation accuracy using deep learning	Technique to be used
5941		0.9966	deep learning
743		0.9864	deep learning
74		0.9202	deep learning
38	0.8429	0.8429	Transfer/deep learning
7	0.6605		transfer learning
1	0.5593		transfer learning

VI. CONCLUSION

This paper has proposed a methodology that adapts existing partial discharge (PD) diagnostic models, originally developed for atmospheric conditions, to high-pressure dry-air environments by leveraging transfer learning techniques. We evaluate the predictive performance under varying degrees of data scarcity in high-pressure dry-air switchgears and analyze the transition point at which deep learning methods become less effective compared to transfer learning as the dataset expands.

Future work will focus on addressing classification challenges under severe class imbalance in defect-type samples. Additionally, we aim to develop more rigorous approaches for determining transition points between deep learning and transfer learning methods, incorporating factors such as sample similarity, data distribution, and the accuracy requirements of specific applications.

REFERENCES

[1] Wong Jee Keen Raymond, Hazlee Azil Illias, Ab Halim Abu Bakar, Hazlie Mokhlis,"Partial discharge classifications: Review of recent progress", Measurement, vol. 68, pp. 164-181, 2015.

[2] High-voltage test techniques – Partial discharge measurements, IEC 60270:2000, 2000.

[3] S. Lu, H. Chai, A. Sahoo and B. T. Phung, "Condition Monitoring Based on Partial Discharge Diagnostics Using Machine Learning Methods: A Comprehensive State-of-the-Art Review," in IEEE Transactions on Dielectrics and Electrical Insulation, vol. 27, no. 6, pp. 1861-1888, December 2020.

[4] R. Bruetsch, M. Tari, K. Froehlich, T. Weiers and R. Vogelsang, "High Voltage Insulation Failure Mechanisms," Conference Record of the 2008 IEEE International Symposium on Electrical Insulation, Vancouver, BC, Canada, 2008, pp. 162-165.

[5] L. Li, C. Wang and M. Yang, "Research on Partial Discharge Detection and Evaluation Method for High Voltage Switchgear Based on AHP-Fuzzy Comprehensive Evaluation," 2024 7th Asia Conference on Energy and Electrical Engineering (ACEEE), Chengdu, China, 2024, pp. 144-149.

[6] T. Hoshino, S. Maruyama, K. Nojima and M. Hanai, "A unique sensitivity verification combined with real-time partial-discharge identification method," in IEEE Transactions on Power Delivery, vol. 20, no. 3, pp. 1890-1896, July 2005.

[7] Q. Khan, S. S. Refaat, H. Abu-Rub and H. A. Toliyat, "Partial discharge detection and diagnosis in gas insulated switchgear: State of the art," in *IEEE Electrical Insulation Magazine*, vol. 35, no. 4, pp. 16-33, July-Aug. 2019.

[8] L. Duan, J. Hu, G. Zhao, K. Chen, J. He and S. X. Wang, "Identification of Partial Discharge Defects Based on Deep Learning Method," in IEEE Transactions on Power Delivery, vol. 34, no. 4, pp. 1557-1568, Aug. 2019.

[9] N. Pattanadech, P. Nimsanong, S. Potivejkul, P. Yuthagowith and S. Polmai, "Partial discharge classification using probabilistic neural network model," 2015 18th International Conference on Electrical Machines and Systems (ICEMS), Pattaya, Thailand, 2015, pp. 1176-1180.

[10] T. Cheypoca, W. Promphanich, A. Y. Thway, A. P. Hankae, S. Jeenmuang and N. Pattanadech, "Partial Discharge Classification with Transformer Neural Networks," 2024 IEEE 14th International Conference on the Properties and Applications of Dielectric Materials (ICPADM), Phuket, Thailand, 2024, pp. 93-96.

[11] lS. Mantach, H. Janani, A. Ashraf and B. Kordi, "Classification of Partial Discharge Signals Using 1D Convolutional Neural Networks," 2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), ON, Canada, 2021, pp. 1-5

[12] Yuwei, Fu, et al. "Partial discharge pattern recognition method based on transfer learning and DenseNet model." IEEE transactions on dielectrics and electrical insulation 30.3 (2023): 1240-124.

[13] Z. Tang and Z. Cao, "Application of Convolutional Neural Network Transfer Learning in Partial Discharge Pattern Recognition," 2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE), Beijing, China, 2020, pp. 1-4.

[14] N. F. A. Aziz, L. Hao and P. L. Lewin, "Analysis of Partial Discharge Measurement Data Using a Support Vector Machine,"2007 5th Student Conference on Research and Development, Selangor, Malaysia, 2007, pp. 1-6.

[15] N. He, H. Liu, M. Qian, W. Miao, K. Liu and L. Wu, "Gas-insulated Switchgear Discharge Pattern Recognition Based on KPCA and Multilayer Neural Network," 2021 6th International Conference on Communication, Image and Signal Processing (CCISP), Chengdu, China, 2021, pp. 208-213.

[16] M. H. H. A. Majid, N. M. Samsuddin, A. S. A. Rahman, N. H. Nik Ali and M. H. F. Rahiman, "Online Partial Discharge Monitoring In Gas Insulated Switchgear (GIS) and Air Insulated Switchgear (AIS)," 2024 14th International Conference on System Engineering and Technology (ICSET), Bandung, Indonesia, 2024, pp. 118-122.

[17] Zawislak, Remy, Greiff, Marcus, Kim, Kyeong Jin, Berntorp, Karl, Di Cairano, Stefano, Konishi, Mao, Parsons, Kieran, Orlik, Philip V., Sato, Yuki, "GNSS Multipath Detection Aided by Unsupervised Domain Adaptation", Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022), Denver, Colorado, September 2022, pp. 2127-2137.

[18] Ilya Loshchilov and Frank Hutter, "Decoupled Weight Decay Regularization", Proceedings of the 7th International Conference on Learning Representations (ICLR 2019), New Orleans, Louisiana, May 2019.